ACHIEVEMENT STANDARD 90641

TIME SERIES

NOTES FOR TEACHERS

This document has been prepared to provide some background information relevant to, and assist teachers with interpreting the requirements of, the achievement standard 90641

ACHIEVEMENT STANDARD 90641 TIME SERIES

NOTES FOR TEACHERS

This document has been prepared to provide some background information relevant to, and to assist teachers with interpreting the requirements of, the achievement standard 90641 *Determine the trend for time series data*.

GENERAL

Some background information should be given to students to enable relevance and usefulness of the forecast (one of the aspects listed for Excellence) to be discussed without having to resort to uninformed speculation. Comments made should be consistent with the background information.

If any of the possible aspects for Excellence that are listed in the standard are not relevant to a particular situation, reference to them should be omitted from the assessment task.

Before embarking on any analysis, students need to develop an understanding of the data by thinking about the variables and what they mean, and how the data may have been collected and for what purpose. This is another reason for providing some background information.

In the first instance, a linear trend line will need to be fitted to the smoothed series so that the gradient (trend) can be interpreted (a requirement for Achievement).

Comments made by students for Excellence need to be justified – e.g. for any proposed improvements to a model an explanation should be given why/how the proposed model is an improvement.

The following headings relate to specifications of the achievement standard for Excellence.

EXCELLENCE

1 Relevance and usefulness of forecast

Discussion could include who would or could (realistically) make use of the forecast and for what purpose, and possibly the extent to which the data and forecast would enable them to do this.

2 Features of the time series data

Features discussed could include

- the nature of the trend
- how well the trend line fits the moving means and the nature of the variations of the moving means about the trend line
- the variability of the individual seasonal effects for a particular season and how these affect the time series graph
- any special features such as outliers.

Discussion needs to be more than just comment on the appearance of the time series graph (which is no more than is expected for NCEA level 1) and needs to relate to the analysis. As an example, the additive time series model is taken to be

$$Y_t = T_t + S_t + E_t$$

where each observation Y_t for any period t consists of a trend component T_t, a seasonal component S_t and a random error component E_t; the model assumes that the seasonal effect for a particular season is constant, so the individual seasonal effects should not vary too much; significant variations in the individual seasonal effects for a particular season will have implications for the analysis.

3 Appropriateness of the model

Discussion needs to refer to the model and not to aspects such as the data collection process.

Comment should be made about how well a linear model fits (as noted previously, a linear model needs to be fitted initially). The fit of the trend line should be based primarily on a visual inspection of the graph, and could include reference to R^2 (but note that with time series the smoothing of the data means that R^2 could be expected to be high).

4 Improvements to the model

Discussion could include the possibility of a non-linear model or a piece-wise model for the trend, or the possibility of a multiplicative time series model. Note that an improved model should be based primarily on a visual inspection of the fit of the trend function to the smoothed series. It should not be based on an increased value of R^2 - in particular, R^2 values should not be the basis for comparing different forms of model - refer to comments on R^2 in the section Notes on *R*-squared and *r* (pages 9 – 12) in the notes for achievement standard 90645.

Ideally, the sum of the estimated seasonal effects should be 0, but in practice it is unlikely to be the case. Corrections to the estimated seasonal effects are sometimes made to compensate for this. Comment could be made on the sum of the estimated seasonal effects and how it could be used to improve the model.

It is unlikely that "having more data" would offer an improvement for time series analysis since the additional data would usually be more historic than the data used and circumstances may be different with older data. However, more data would provide more values to use in calculating the estimated seasonal effects, and more values to consider the variability in the individual effects.

5 Limitations of the analysis

Discussion must relate to the analysis and not to aspects such as the data collection process.

Discussion could include:

- reference to the (fixed) weightings used in moving average smoothing models (other models give greater weight to the most recent data on the basis that these data are more important in determining subsequent values than earlier data)

- the danger of extrapolating too far beyond the observed data (with the need to ensure that conditions remain constant)
- the danger of extrapolating with a polynomial model if it is proposed as an improvement (see the comment on page 11 of the notes for achievement standard 90645)
- the gap at the end (and start) of the series of centred moving means - for quarterly data, this means that there is no value for either of the last two quarters that contribute to the determination of the trend line, with the missing values being in periods immediately preceding periods where predictions are to be made; other methods of time series analysis overcome this limitation.

Teachers may wish to distinguish between the *validity* of a forecast and the *reliability* of a forecast.

- Validity refers to the process by which a forecast has been obtained. A least squares regression model fitted to moving means (centred where necessary) provides a valid estimate of trend. Comment could be made on the process used to estimate seasonal effects.
- Reliability refers to how good a forecast is. Comment could be made on how good the estimate of the trend component of a forecast is and how good the estimated seasonal effect is.

This aspect is closely related to improvements to the model, and the two could be combined together in the assessment task.

6 Seasonally adjusted data

Seasonally adjusted (deseasonalised) data are obtained by subtracting from each observation the estimated seasonal effect (the mean of the individual seasonal effects) for the corresponding season. Discussion should include calculation and explanation (interpretation) of the seasonally adjusted data, using sufficient seasons to show that their use is understood.

Seasonally adjusted data provide a comparison over consecutive seasons of movements in the trend (including random error effects – see comment following); these movements are masked (smoothed out) by the trend line. Seasonally adjusted data allow the question "Is an increase or a decrease in the observed variable likely to be a real increase or decrease?" to be addressed. An increase or decrease in data that have not been seasonally adjusted may be (primarily) due to seasonal effects. To illustrate this further, consider the following hypothetical unemployment data for a particular region over successive quarters (unemployment is seasonal and so is a good example to use).

Year	Quarter	Number of registered unemployed
1	M	2500
	J	2600
	S	2700
	D	2650
2	M	2550
	J	2550
	S	2550
	D	2550

Reading down the last column, it appears that unemployment is rising for the first three quarters, then falling for two quarters, and is then static over the final four quarters. However, what we may be seeing is the influence of the seasonal effects (and/or random error effects). To illustrate this, suppose that the estimated seasonal effects for each quarter are as follows:

Quarter	Estimate of seasonal effect
M	150
J	300
S	250
D	200

Then the seasonally adjusted data are as shown in the following table:

Year	Quarter	No of unemployed	Seasonally adjusted data
1	M	2500	2500 – 150 = 2350
	J	2600	2600 – 300 = 2300
	S	2700	2700 – 250 = 2450
	D	2650	2650 – 200 = 2450
2	M	2550	2550 – 150 = 2400
	J	2550	2550 – 300 = 2250
	S	2550	2550 – 250 = 2300
	D	2550	2550 – 200 = 2350

The seasonally adjusted values show that unemployment falls for one quarter, then rises for one quarter, remains the same for the next quarter, then falls for two quarters, and then rises again over the final two quarters. This is a different picture from that previously. Removing the seasonal effects has allowed short term variations in the trend to be seen - what is really happening (such as the effects of structural economic factors) without the "contamination" of seasonal influences.

It should be noted that the influence of some random error effects could still be present in the seasonally adjusted data. As noted above, the additive time series model is taken to be

$$Y_t = T_t + S_t + E_t$$

so that $$Y_t - S_t = T_t + E_t$$

and subtracting the (estimated) seasonal effect from an observation leaves the trend and random error effect.

Seasonally adjusted series are important in trying to identify economic turning points, where the economy changes direction. These are likely to be seen first by looking at seasonally adjusted macroeconomic data such as national unemployment or retail sales data. However, seasonal patterns are never totally regular, and a seasonally adjusted series may only be a guide to the underlying changes in the trend. In addition, the seasonal effects are estimated values (the means of the individual seasonal effects for each season) and this may have some affect on the seasonally adjusted data. Hence the seasonally adjusted data must be interpreted carefully with these aspects in mind (particularly when the values are numerically close).

It should be noted that the term *trend* is used in slightly different ways in relation to time series.

- As used in relation to the model, trend refers to the overall or long-term tendency.
- Trend also refers to the function from which the trend components of forecasts are obtained.
- As used in relation to forecasts, trend refers to the numerical value of the trend component of the forecast.

7 Comparison with related time series

This is only relevant if related time series data are available.

8 Development and interpretation of an index number series

This is only relevant if students understand the nature of index numbers.